New Applications of the Least Squares Method in Spectroscopic Characterization of Thin Organic Films

Keiichi Ikegami

Nanotechnology Research Institute, National Institute of Advanced Industrial Science and Technology
AIST Tsukuba Central-2, 1-1-1 Umezono, Tsukuba 305-8568, Japan
Fax: 81-29-861-5400, e-mail: k.ikegami@aist.go.jp

In these years the principal component analysis method has been used widely in spectroscopy, while the use of the least squares method (LSM) has been limited. One of the reasons caused this situation is that the classic application of LSM to the Lambert-Beer law is unreliable unless the concentrations of chemical species are exactly known. However, in principle, the concentrations of chemical species can be treated as unknown parameters that should be estimated by LSM. In this context, a new procedure based on LSM, optimum difference spectrum (DOS) method, has been proposed. The DOS method takes one of the advantages of spectroscopy, i.e., the difference between the number of data and that of unknowns can be very large. Taking the same advantage, the thickness and the wavelength-dependent anisotropic optical constants of organic thin films can be estimated simultaneously from their spectroscopic data taken by a conventional spectrometer.
Key words: Least squares method, difference spectrum, spectrum analysis, FT-IR, electronic absorption

1. INTRODUCTION

Spectroscopy is a powerful tool for characterizing materials because it provides much information about the samples. In spectroscopic studies, therefore, such information should be extracted as much as possible from the experimentally obtained spectra.

For example, when the mixtures of several chemical species are spectroscopically studied under the condition that the pure materials are not available, separation of the measured spectra into the spectrum of each species, and/or estimation of the concentration of each species are essential. Difference spectrum method [1] is a simple and useful method for spectrum analysis in this case. It is commonly used to extract the spectrum of an organic species from the spectrum observed for its solution. It was also applied to a Langmuir-Blodgett (LB) film consisting of two components, one of which shows in-plane anisotropy and the other does not, and the spectrum originating from the former component was extracted [2]. However, from the statistical viewpoint, information distributed in more than three spectra cannot appropriately be extracted by this method.

On the contrary, the principal component analysis (PCA) method is based on statistics [3] and can appropriately extract information from a number of spectra. More importantly, it does not require any information other than the spectra, because it deconvolute the *k* observed spectra into *k* mathematically orthogonal components without referring to physical models [3], where *k* is the number of the observed spectra and not relevant with the number of the chemical species in the samples. This deconvolution can always be done

regardless of the quality of the experiments, i.e., PCA is a robust method. These are greatly advantageous and accordingly PCA is widely used in spectroscopy in these days. However, the connection between the chemical species and the deconvoluted components is unclear and additional complicated calculation is needed to relate them with each other.

Least squares method (LSM) may be intermediate between the difference spectrum method and PCA, i.e., it is based both on statistics and closely connected to physical models. Consequently, the results of LSM are directly related to physical quantities, and additional information other than the spectra is indispensable for constructing the physical models. These are advantage and disadvantage of LSM and the latter has limited the application of LSM. Actually, the classic application of LSM to spectrum separation is unreliable except when the exact concentrations of the constituents are known [3].

Incidentally, a quantity measured in spectroscopy (e.g., absorbance) of the *i*-th sample at the *j*-th wavelength may be a function of *i* and *j* written as

$$A_{ij} = f(\boldsymbol{a}_i, \boldsymbol{b}_j; i, j) , \qquad (1)$$

where $\boldsymbol{a}_i$ is a vector containing the parameters depending only on the sample and $\boldsymbol{b}_j$ is that containing the parameters depending only on the wavelength. These parameters may be unknowns that should be estimated through spectrum analysis. Let us suppose that the spectra of *k* samples are obtained for a wavelength range containing *l* points for analogue-to-digital conversion. In

this case, $n = kl$ data are recorded, while the number of the parameter, m, is $O(k) + O(l)$. If k and l are sufficiently large, the difference of $n - m$ can be very large. This is one of the advantages of spectroscopy and can diminish the disadvantage of LSM described above. In this paper, new applications of LSM in spectroscopy based on this advantage are proposed and demonstrated.

2. BASIC FRAMEWORK

Here, it is supposed that the physical model describing the relationship between a set of quantities X and a quantity Y is given as a function f including unknown parameters of $a_1, a_2, ..., $ and a_m,

$$Y = f(a_1, a_2, \cdots a_m; X), \tag{2}$$

and experiments provide data of (x_i, y_i), where $x_i \in X$, $y_i \in Y$, and $i = 1, 2, ..., n$. The index i may stand for samples, time, or anything that distinguishes the measurements. If the experiments are ideal, i.e., both x_i and y_i are measured without uncertainty, and if $n \geq m$, the unknown parameters can be obtained exactly by solving simultaneous equations of

$$y_i = f(a_1, a_2, \cdots a_m; x_i). \tag{3}$$

In general, uncertainty is an inevitable element of experiments. However, in some cases, the uncertainty of x_i is negligibly smaller than that of y_i. In those cases, eq. (3) can be rewritten as

$$\tilde{y}_i = f(a_1, a_2, \cdots a_m; x_i) + e_i, \tag{4}$$

where e_i is the uncertainty of $\tilde{y}_i$ and the tilde shows that the measurement is not ideal. If e_i is impartial, the maximum likelihood estimators of the parameters minimize the residual sum of squares [4],

$$S = \sum_{i=1}^{n} |e_i|^2. \tag{5}$$

For unique determination of the estimators, of course, the condition $n \geq m$ should be fulfilled, and the reliability of the estimation is proportional to $(n - m)$. This method for estimating unknown parameters is known as LSM.

3. OPTIMUM DIFFERENCE SPECTRUM

Let us consider mixed solutions of chemical species P, Q, ... and their electronic absorption spectra. According to the Lambert-Beer law, the absorbance, A_{ij}, of the i-th sample at the j-th wavelength is given as

$$A_{ij} = p_i P_j + q_i Q_j + \cdots, \tag{6}$$

where $p_i, q_i, ...$ $(1 \leq i \leq k)$ are the concentration of P, Q, ...,

and $P_j, Q_j, ...$ $(1 \leq j \leq l)$ are the absorbance of the pure solutions of P, Q, ... with the unit concentration. Here, $P_j, Q_j, ...$ are unknown. If $p_i, q_i, ...$ are exactly known and the measurement of A_{ij} is ideally performed, difference spectra can give the pure spectra $(P_j, Q_j, ...)$. For example, when only two species are mixed, difference spectrum $(A_{1j} - q_1 A_{2j}/q_2)/p_1$ gives P_j. When $p_i, q_i, ...$ are exactly known and the uncertainty of A_{ij}, e_{ij}, is not negligible, the pure spectra can be estimated by classic application of LSM to the obtained data regarding $(p_i, q_i, ...) \in X$. In reality, however, knowing the exact concentrations is not always possible and the classic application of LSM is not powerful enough. In this section, procedures that are based on LSM and can estimate the pure spectra without information about the concentrations are proposed.

3.1 With an isosbestic point

When the recorded spectra exhibit an isosbestic point, it can be concluded that the samples consist of only two components (P and Q) and the sum of the components' concentrations is a constant. Denoting the averages of p_i and q_i as $<p>$ and $<q>$, respectively, p'_i is defined as

$$p'_i = p_i - \langle p \rangle \ \left(= -q_i + \langle q \rangle\right), \tag{7}$$

and then eq. (6) including the uncertainty can be written as

$$\tilde{A}_{ij} = (p'_i / p'_k) D_j + M_j + e_{ij}, \tag{8}$$

where $D_j = p'_k (P_j - Q_j)$ and $M_j = <p>P_j + <q>Q_j$. Noting that $p'_{k-1}/p'_k = -1 - \sum_{i=1}^{k-2}(p'_i/p'_k)$, it is recognized that the independent parameters in this case are p'_i/p'_k with $i = 1, 2, ..., k-2$ and D_j and M_j with $j = 1, 2, ..., l$. The condition of $m = k + 2l - 2 \leq n = kl$ can be satisfied easily and then LSM is able to estimate those parameters by regarding $(i, j) \in X$.

Here, it should be emphasized that D_j is proportional to the difference between the pure spectra and then its maximum likelihood estimator can be regarded as an *Optimum Difference Spectrum* (ODS) in the present problem. Although M_j may be estimated by simply averaging $\tilde{A}_{ij}$ for each j, D_j cannot appropriately be estimated by the classic difference spectrum method and application of LSM is necessary. Thus, the present procedure is named as *optimum difference spectrum* method.

Efficiency of the ODS method in spectrum analysis has been examined by applying it to $k = 10$ simulated spectra (Fig. 1 (a)) that are sums of $p_i P_j$, $q_i Q_j$, and Gaussian noise, where $p_i + q_i$ is fixed and P_j and Q_j are given in Fig. 1 (b) (dashed lines). The mixing ratios are listed in Table I. The initial estimators of M_j and D_j have

been given by the average of $\tilde{A}_{ij}$ and the difference of $\tilde{A}_{1j} - \tilde{A}_{10j}$. Those of p'_i/p'_{10} have been set to be regularly distributed from –1 to +1 for $i = 1$ to 10.

The estimated M_j and D_j are displayed in Fig. 1 (b). Gaussian noise is remarkably reduced in M_j and D_j, proving the noise reduction ability of LSM. (With the simple difference spectrum method, noise is generally enhanced.) The estimators of p'_i/p'_k are compared to their true values in Table I. Good agreement between the estimators and true values of p'_i/p'_k implies that the application of the present procedure is successful.

Unfortunately, p'_k, $<p>$ and $<q>$ cannot be estimated in a mathematical way and the maximum likelihood estimation of the pure spectra P_j and Q_j is impossible. However, if it is possible to assume that the absorption bands have rather simple and structureless lineshapes, the elimination of unnatural structures in linear combinations of M_j and D_j provides physical estimators of P_j and Q_j as

$$P_j \propto M_j + sD_j \quad \text{and} \quad Q_j \propto M_j - tD_j ,\qquad (9)$$

where s and t should be determined by the analyst. Such
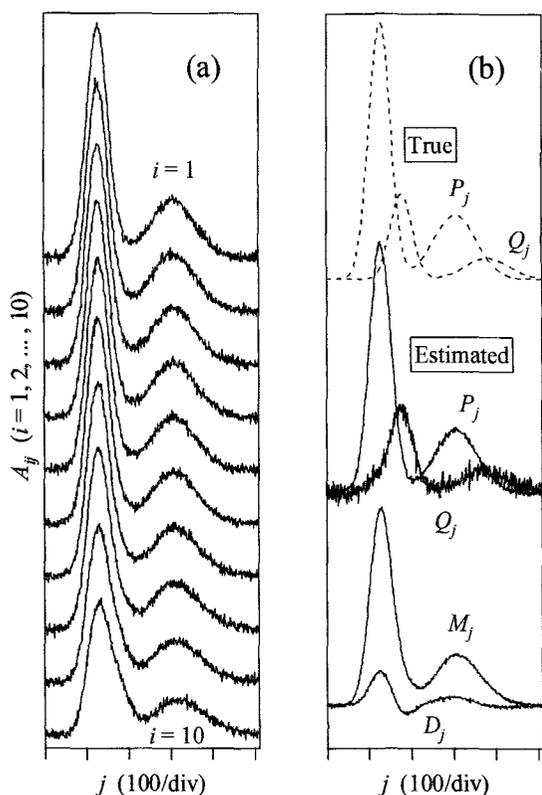

Fig. 1 (a) Simulated spectra that are sums of p_iP_j, q_iQ_j, and Gaussian noise, where $p_i + q_i$ is fixed and P_j and Q_j are represented by the dashed lines in (b). (b) True (---) and estimated (–) P_j and Q_j, and estimated D_j (optimum difference spectrum) and M_j.

estimators of P_j and Q_j in the present simulation are displayed in Fig. 1 (b) and satisfactorily coincide with the true P_j and Q_j, respectively.

Furthermore, the relative concentrations can be estimated by using the determined s and t values as

$$\frac{p_i}{p_i + q_i} = \frac{(p'_i/p'_k) + t}{s + t}\qquad (10)$$

The relative concentrations estimated by this method are listed in Table I. They show good agreement with the true values.

No correlation between p_i and i was introduced in the simulation discussed above. In some cases, however, a physical model with a few parameters, $a_1, a_2, ..., a_{k'}$, can relate p_i to i. Such a physical model should be incorporated into the ODS method and the maximum likelihood estimators of $a_1, a_2, ..., a_{k'}$ may be determined when $k' \leq k - 2$. In those cases, the maximum likelihood estimators of P_j and Q_j will be determined. (The adjustment of s and t in eq. (9) is not necessary.) For example, in a simple decay case, $p_i = a_1\exp(-i/a_2)$ can be supposed and all of a_2, a_1P_j, and a_1Q_j, will be determined by the ODS method.

3.2 Two component system

When $p_i + q_i$ is not constant, no isosbestic point is observed even for two component systems, and the number of independent parameters is increased to $m = 2k + 2l - 4$. Instead of eq. (8), the next representation is useful in this case:

$$\tilde{A}_{ij} = p''_iP'_j + q''_iQ'_j + e_{ij} ,\qquad (11)$$

where

Table I Comparison of the true and estimated values of p'_i/p'_k and $p_i/(p_i + q_i)$ in the trial application of optimum difference spectrum method to the simulated data shown in Fig. 1.

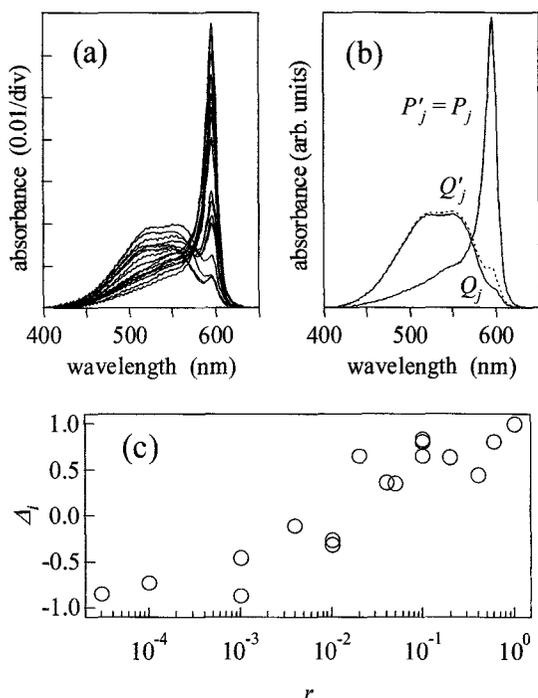| i	true values		estimated values	
	$p_i/(p_i + q_i)$	p'_i/p'_k	p'_i/p'_k	$p_i/(p_i + q_i)$
1	0.5	–1.96	–1.97	0.48
2	0.6	–1.22	–1.22	0.59
3	0.7	–0.48	–0.49	0.70
4	0.75	–0.11	–0.11	0.76
5	0.8	0.26	0.26	0.82
6	0.82	0.41	0.40	0.84
7	0.84	0.56	0.55	0.86
8	0.86	0.70	0.71	0.88
9	0.88	0.85	0.86	0.90
10	0.9	1.00	1.00	0.93

Fig. 2 (a) Electronic absorption spectra observed for Langmuir films of $1{:}r$ mixtures of DSe and arachidic acid $(0 \leq r \leq 1)$. (b) ODS's and pure spectra estimated from the data in (a). (c) The factor Δ_i as a function of r.

$$p_i'' = \frac{p_i q_2 - p_2 q_i}{p_1 q_2 - p_2 q_1}, \quad q_i'' = \frac{p_1 q_i - p_i q_1}{p_1 q_2 - p_2 q_1} \quad (12)$$

$$P_j' = p_1 P_j + q_1 Q_j, \text{ and } \quad Q_j' = p_2 P_j + q_2 Q_j \quad (13)$$

are newly defined DOS's for $i = 3, 4, \ldots k$ and $j = 1, 2, \ldots l$. (In this case, the spectra seem to correspond to the smallest p_i and q_i may be numbered as 2 and 1, respectively, for convenience.)

The pure spectra P_j and Q_j can physically be estimated by making linear combinations of P_j' and Q_j' as in the cases with isosbestic points as

$$P_j \propto P_j' - s'Q_j' \text{ and } Q_j \propto Q_j' - t'P_j', \quad (14)$$

where s' and t' are adjusted to minimize the unnatural structures in these linear combinations. This is why P_j' and Q_j' can be regarded as ODS's in this case. Furthermore, the s' and t' can be used to obtain a factor Δ_i that indicates the predominance of the species P or Q as

$$\Delta_i = \frac{(1-s')p_i'' + (t'-1)q_i''}{(1+s')p_i'' + (t'+1)q_i''}. \quad (15)$$

The Δ_i value of $+1$ (-1) suggests that the i-th sample contains P (Q) only, and that of zero suggests that the sample is the 1:1 mixture.

This procedure was applied to the *in-situ* electronic

absorption spectra observed for Langmuir films of the mixtures of a merocyanine derivative, DSe, and arachidic acid with the molar ratio of $1{:}r$ $(0 \leq r \leq 1)$ [5,6]. When these films are prepared on a pure-water subphase, the formation of the J-aggregates of DSe depends on r. In these films, the DSe molecules that do not form J-aggregates form another type of dye aggregate, named "D-aggregate". Namely, the system consists of two optically significant components, J- and D-aggregates, and the concentration of the J-aggregated dye, whose absorption band is narrowed and located around 600 nm, and that of D-aggregated dye, whose absorption band is a doublet around 530 nm, vary from sample to sample (Fig. 2 (a)). The ODS method provided P_j' and Q_j' mathematically and then P_j and Q_j physically (Fig. 2 (b)). Here, P_j and Q_j can be regarded as the pure spectra of the J- and D-aggregates. More importantly, the relationship between the factor Δ_i and r was given (Fig. 2 (c)). From the fact that $\Delta_i \approx 0$ around $r = 10^{-2}$, it was concluded that one arachidic acid molecule converts 10^2 DSe molecules through the acceleration of nucleation.

3.3 ODS for joined spectra

In the procedure described in the sections 3.1 and 3.2, the index j is regarded as a part of X, but the real wavelength values are not referred to. This fact allows us to join several different spectra observed for each sample i into one "joined spectrum". For example, the electronic absorption spectra taken with different angle of incidence and optical polarization can be joined. Optical spectra recorded by different apparatus, such as those recorded by a dispersion-type ultraviolet (UV) spectrometer and a Fourier-transform infrared (IR) spectrometer, can be joined, too. Rather, any quantity keeping a linear relationship with the concentration (p_i and q_i) can be included in a joined spectrum $\tilde{A}_{ij}$ for each sample i and the corresponding ODS's can be estimated.

Let us suppose that the UV and IR spectra of two-component samples are observed. If the UV and IR spectra are analyzed separately, the estimators of p_i and q_i obtained from UV spectra and those from IR spectra will not exactly coincide with each other. The analyst may adopt the averaged values for the concentrations, but they are not consistent with the "pure spectra" estimated in each analysis. On the contrary, if the UV and IR spectra are joined and the ODS method is applied to the joined spectra, the consistency concerning the concentrations is automatically assured. Furthermore, joining the UV and IR spectra may remove the difficulty in estimating s and t (s' and t') for complicated IR spectra. The estimation of s and t is based on the assumption that the pure spectra are rather simple and structureless. This assumption is not appropriate for IR spectra in general. However, when the ODS method is applied to the joined spectra, s and t can be estimated by eliminating unnatural structures in the
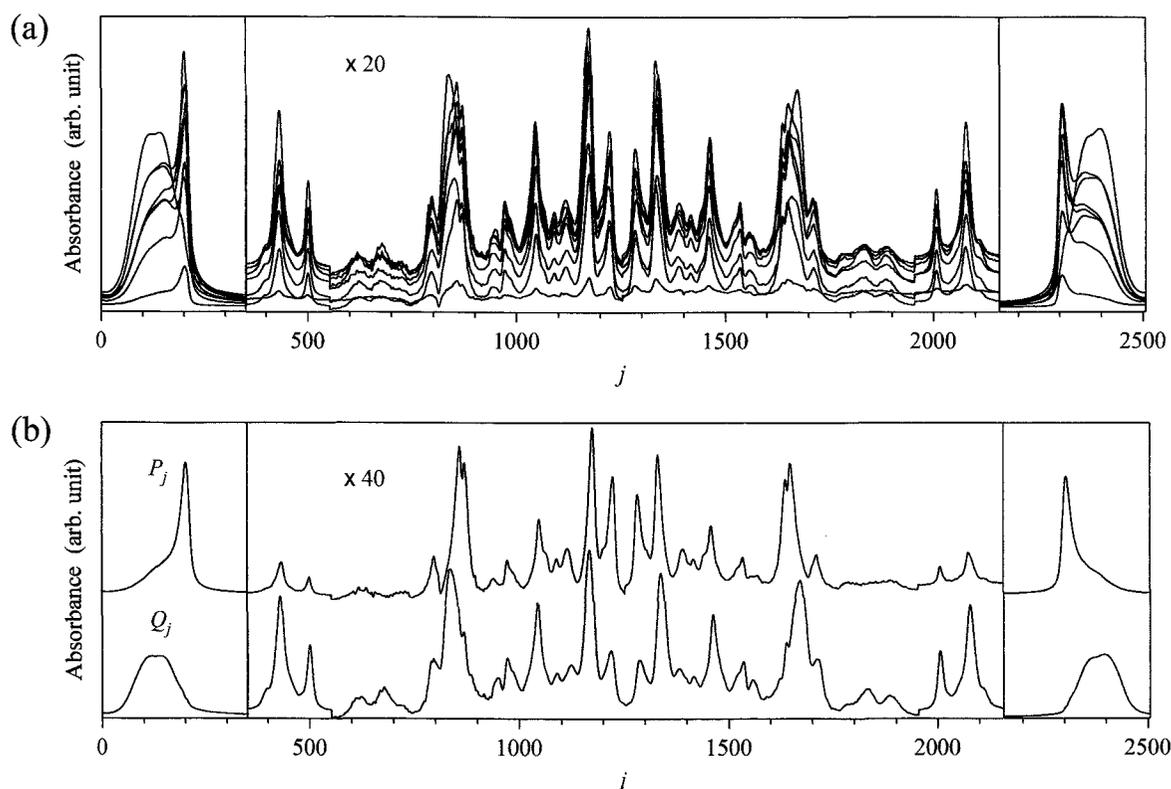
Fig. 3 (a)Joined spectra containing the UV and IR spectra of LB films of 6MeDS observed for s- and p- polarized light with the angle of incidence of 45°. (b) Pure spectra estimated from the data in (a).

Table II Correspondence between the index j and the real wavelength (λ_j) or wavenumber (ν_j) and polarization of light in the joined spectra shown in Fig. 3.

j	λ_j (nm) or ν_j (cm^{-1})	polarization
$1 \leq j \leq 351$	$\lambda_j = j + 399$	s
$352 \leq j \leq 552$	$\nu_j = -j + 3352$	s
$553 \leq j \leq 1253$	$\nu_j = -j + 2353$	s
$1254 \leq j \leq 1954$	$\nu_j = j - 154$	p
$1955 \leq j \leq 2155$	$\nu_j = j + 845$	p
$2156 \leq j \leq 2506$	$\lambda_j = -j + 2906$	p

UV region and the pure spectra in the IR region are automatically obtained. (The latter advantage of joined spectra was already proved in the spectrum analysis based on the simple difference spectrum method [2].)

The ODS method was applied to joined spectra for characterizing the J-aggregates of another merocyanine dye derivative, 6MeDS [7]. This dye forms J-aggregates in its pure Langmuir films prepared on a pure-water subphase. However, these J-aggregates are not very stable and after transferred onto CaF$_2$ substrates the films consist of two components, i.e., J-aggregated dye and D-aggregated dye, and their concentrations randomly vary from sample to sample. (The transfer ratio varies from sample to sample, too.) The UV and IR spectra of the transferred films were recorded for s- and p-polarized

light with the angle of incidence of 45°, and then joined (Fig. 3 (a)). Here, the correspondence between the index j and the wavelength (or wavenumber) is given in Table II. The ODS method applied to the joined spectra well separated the IR spectrum of the J-aggregates (P_j) and that of the D-aggregates (Q_j), as shown in Fig. 3 (b). Following remarkable differences between these spectra were detected. (i) The absorption band due to the free keto group is observed around 1678 cm^{-1} for the D-aggregates, but not for the J-aggregates. (ii) The absorption band due to the central conjugated system and that due to the C=S group, which are observed around 1516 and 1185 cm^{-1}, respectively, for the D-aggregates, are observed around 1496 and 1178 cm^{-1}, respectively, for the J-aggregates. (iii) The bands mentioned in (ii) are narrowed in the J-aggregates. Based on these observations, the structure and the formation mechanism of the J-aggregates of 6MeDS could be discussed.

Extension of the ODS method to multi-component system is straightforward. The increase in the number of the components increases the number of unknowns and consequently the convergence of LSM may become less stable. However, if some physical model describing the relationship between the index i and the concentrations of the components can be assumed, the incorporation of that physical model may greatly empower the ODS method. Namely, the convergence of LSM may become stable and

maximum likelihood estimation of pure spectra may become possible.

4. DISPERSION OF OPTICAL CONSTANTS

In this section, a new procedure applicable to a sub 100 nm-thick organic film deposited on a transparent substrate for simultaneous estimation of its thickness and wavelength-dependent and anisotropic optical constant. In this case, transmission or reflection spectra of one single sample, as well as those of the bare substrate, are measured under several conditions. To adapt to this problem, A is redefined as

$$A = \log\left(T_{sub}/T_{smpl}\right) \quad \text{or} \quad \log\left(R_{sub}/R_{smpl}\right) \tag{16}$$

in this section, where T_{sub} and R_{sub} are the transmittance and reflectance of the bare substrate, respectively, and T_{smpl} and R_{smpl} are those of the sample, respectively. This A can be related to the thin film's thickness, d, and its optical dielectric tensor, $\tilde{\varepsilon}$, and an integer indexing the experimental conditions, i, as

$$A_i = f(d, \tilde{\varepsilon}; i) . \tag{17}$$

Here, the index i (= 1, 2, ... k) determines the type of measurement (transmission or reflection), the angle of incidence, the azimuth of the plane of incidence, and the polarization of light (p or s). The actual formula of f in eq. (17) can be deduced based on Fresnel theory of reflection and that of double refraction and has been given Refs [8] and [9]. (It is very complicated and not explicitly given in this paper.)

When the principal axes are known and the wavelength is fixed, the number of unknowns in eq. (16) is seven. Therefore, the unknowns can be determined, if more than seven data sets without uncertainty are obtained. However, the dependence of A_i on d and that on $\tilde{\varepsilon}$ are so closely correlated that the uncertainty of the experimentally recorded A_i is usually too large to determine d and $\tilde{\varepsilon}$ when d is smaller than the wavelength. As a consequence of this difficulty, d and $\tilde{\varepsilon}$ of thin organic films are usually determined from ellipsometric data, which are recorded by costly apparatuses.

The difficulty caused by the uncertainty can be reduced by applying LSM to the data recorded in a wide wavelength range,

$$\tilde{A}_{ij} = f(d, \tilde{\varepsilon}_j; i, j) + e_{ij} , \tag{18}$$

where the integer j (= 1, 2, ... l) indexes the wavelength. Although the number of unknowns increases to $m = 6l + 1$, the increase in the number of data $n = kl$ is more rapid when $k \geq 7$. Consequently, the large difference of $n - m$

may enable the maximum likelihood estimation of d and $\tilde{\varepsilon}_j$.

The proposed procedure was applied to the UV–visible–near IR (280 – 1500 nm) transmission and reflection spectra of LB films of cadmium arachidate recorded by a conventional spectrometer [8]. Both d and $\tilde{\varepsilon}_j$ could simultaneously be estimated for 25- and 37-layer samples, although such a simultaneous estimation for a 13-layer sample was unsuccessful. The real part of the estimated in-plane optical constant is ca. 2.45 in the UV region and decreases to ca. 2.25 in the near IR region, showing a reasonable dispersion behavior, while its imaginary part is less than 5×10^{-3}. Moreover, the d value estimated for the 25-layer sample (68.7 nm) coincides with that deduced from X-ray diffraction within 1%.

5. CONCLUSIONS

New applications of LSM in spectroscopic studies are proposed. Their efficiencies were examined by applying them to simulated and experimentally obtained spectral data and then were proved. These new applications of LSM make full use of one of the characteristics of spectroscopy, i.e., the difference in the numbers of data and unknowns can be very large. The Optimum Difference Spectrum (ODS) method is useful in spectrum analysis, especially when it is applied to joined spectra that contain several kinds of spectra (visible, infrared, electron spin resonance, ...). Another application of LSM enables simultaneous determination of the thickness and the wavelength-dependent anisotropic optical constant of thin organic films deposited on transparent substrate.

REFERENCES

[1] A. Ishitani, in "Fourier-transform infrared spectroscopy" J. Hiraishi, ed. (Gakkai Shuppan Center, Tokyo, 1985) Chap. 4, pp. 149-157.

[2] K. Ikegami, C. Mingotaud and M. Lan, *J. Phys. Chem. B*, **103**, 11261-11268 (1999).

[3] T. Hasegawa, *Bunko Kenkyu*, **49**, 175-185 (2000).

[4] M. Homma and N. Kasugaya, "Dimensional Analysis / Least Squares Method and Experimental Formula" (Corona, Tokyo, 1957) pp. 241-262

[5] K. Ikegami, *Jpn. J. Appl. Phys.* **41** 5444 –5448 (2002).

[6] K. Ikegami, *Trans. Mater. Res. Soc. Jpn.* **28** 31-34 (2003).

[7] K. Ikegami, *J. Chem. Phys.* **121** 2337-2347 (2004).

[8] K. Ikegami, *Thin Solid Films, in press.*

[9] K. Ikegami, *Bull. Electrotechn. Lab.* **61** 247-252 (1997).